DOCUMENT RESUME

ED 440 118 TM 030 722

AUTHOR Ushijima, Teri; Edwards, Jenny
TITLE Program Evaluation: Showing That You Are Making a Difference.
PUB DATE 1998-12-09
NOTE 40p.; Paper presented at the Meeting of the Staff Development Council (Washington, DC, December 9, 1998).
PUB TYPE Guides - Non-Classroom (055)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cost Effectiveness; Data Analysis; *Data Collection; *Evaluation Methods; Hypothesis Testing; *Information Dissemination; *Program Evaluation; *Research Methodology

ABSTRACT

This guide to program evaluation presents questions that illuminate aspects of the evaluation process. Issues to consider in program evaluation are: (1) the intended results of the program; (2) hypotheses about the possible results of the program; (3) the types of data that will indicate achievement of the intended results and the variables that will be measured; (4) the measures that will indicate achievement of the intended results of the program; (5) data sources; (6) research design; (7) the timeline for gathering and analyzing data; (8) data collection; (9) data analysis; (10) data interpretation; (11) adjustment to the project based on findings; (12) communication of the results; and (13) the approximate cost of the evaluation. (Contains 45 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 440 118

Program Evaluation: Showing That You Are Making a Difference

National Staff Development Council
December 9, 1998
Washington, D. C.

Presented by:

Jenny Edwards, Ph.D.
3774 Mountainside Trail
Evergreen, Colorado 80439
(303) 674-0645 (o)
(303) 674-5007 (FAX)
jedwards@fielding.edu
&
Teri Ushijima, Ed.D.
98-2060 Kikala Street
Aiea, HI 96701
(808) 626-3144 (o)
(808) 487-8423 (FAX)
teri@lava.net

TM030722

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
J. Edwards
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Table of Contents

Questions to Ask When Evaluating a Program 3

Intended Results 4

1. What are the intended results of the program? 5

Hypotheses About the Possible Results 6

2. What hypotheses do you have about the possible results of the program? 8

Types of Data - Variables - Measures 9

Quantitative Data 9

Qualitative Data 12

Terms 14

Sources of Instruments and Evaluation Strategies 15

Issues to Consider 16

3. What types of data will indicate achievement of the intended results of the program? What are the variables? . . 17

4. What measures will indicate achievement of the intended results of the program? 18

Sources of Data / Participants 19

5. Who will provide data to indicate success of the program? 20

Types of Research Designs 21

6. What type(s) of research design(s) will be used? 22

7. What timeline will be used for gathering and analyzing the data? 23

8. How will data be collected? 24

Data Analysis 25

9. How will data be analyzed? 26

10. How will the team interpret the data? 27

11. How will the team adjust the project based on findings?28

Communication of Findings 29

12. To whom will the results be communicated? How will they be communicated? 30

13. What will be the approximate cost of the evaluation? . . 31

Bibliography 32

Questions to Ask When Evaluating a Program

1 What are the intended results of the program?

2. What hypotheses do you have about the possible results of the program?

3. What types of data will indicate achievement of the intended results of the program? What are the variables?

4. What measures will indicate achievement of the intended results of the program?

5. Who will provide data to indicate success of the program?

6. What type(s) of research design(s) will be used?

7. What timeline will be used for gathering and analyzing the data?

8. How will data be collected?

9. How will data be analyzed?

10. How will the team interpret the data?

11. How will the team adjust the program based on findings?

12. To whom will the results be communicated? How will they be communicated?

13. What will be the approximate cost of the evaluation?

Intended Results

- What are the intended results of the program? In order to be most effective, evaluation begins with what we are seeking to accomplish through the program.

- *How to Focus an Evaluation* (Stecher & Davis, 1987) provides detailed information about getting started with an evaluation.

- *How to Assess Program Implementation* (King, Morris, & Fitz-Gibbon, 1987) provides an overview of the evaluation process.

1. What are the intended results of the program?

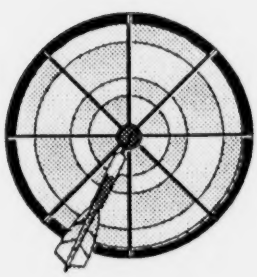
a.

b.

c.

d.

e



Hypotheses About the Possible Results

Four types of claims can be investigated (Ralph & Dwyer, 1988). They include:

- **Academic Achievement** - Changes in Knowledge and Skills of Students - This requires some measure of student learning, examining growth compared with a control group, or other standards.
    - Improved student test scores on standardized instruments
    - Acquisition of factual knowledge
    - Application of knowledge
    - Acquisition of new skills
    - Application of skills
    - Increased speed of learning

- **Improvements in Teachers' Attitudes and Behaviors** - It is important to link teacher changes to educationally important outcomes for students. For example, high teacher efficacy has been found to result in beneficial outcomes for students; therefore, if teachers were to grow in efficacy as a result of a program, it might follow that students would benefit.
    - Improvement on an instrument that has been related to positive outcomes for students
    - Increase in instructional time
    - Changes in instructional methods
    - Increased emphasis on one or more things
    - Higher expectations of students
    - Less time spent in transitions and in getting students' attention

- **Improvements in Students' Attitudes and Behaviors** - It would also be important to link these improvements with educationally beneficial outcomes for students.
    - Increased student attendance
    - Decreased drop-out rate
    - Changes in students' interactions with each other
    - Changes in attitude toward a subject
    - More positive attitudes about self, learning, others, school, etc.
    - Changes in graduation rate
    - Decreased use of drugs

- **Improvements in Instructional Practices and Procedures** - Improvements in this area would also be linked with educationally desirable outcomes for students.
    - Increased numbers of students who are served
    - Reduction in costs
    - Increased circulation of materials
    - Increased used of facilities

2. What hypotheses do you have about the possible results of the program?

a.

b.

c.

d.

e.

Types of Data - Variables - Measures
Quantitative Data

- **Used when:**
  - Specified outcomes can be measured
  - The emphasis is on comparing measurements
  - Results can be captured in numbers
  - The goal is to gather findings that can be generalized to a larger population

- **Attitude Measures**
  - Measure attitudes toward something
  - Can be adapted to a particular audience
  - Can take the form of self-report, reports of others, sociometric procedures, or records (Henerson, Morris, & Fitz-Gibbon, 1987)
  - Generally have a Likert scale, i.e., SA=Strongly Agree, A=Agree, U=Undecided, D=Disagree, and SD=Strongly Disagree
  - Are elaborated on in *How to Measure Attitudes* (Henerson, Morris, & Fitz-Gibbon, 1987) - also includes information about how to develop attitude measures

- **Checklists**
  - Include items to examine
  - Are constructed so that a check mark indicates presence of the attribute, and no check mark indicates absence of the attribute

- **Criterion-Referenced Testing**
  - Assesses a student's performance on knowledge, skills, or standards
  - Includes portfolios, teacher-made tests, performance tests, etc.

- **Demographic Forms**
  - It is helpful to brainstorm all possible variables that might influence responses on other instruments. For example, class size and number of behavior problems might influence a teacher's feelings of efficacy. By gathering these kinds of data, the findings will make more sense. Also, if main effects are not found, the researcher will be able to determine variables that correlate with growth.
  - It is also preferable to gather specific data rather than aggregated data. Instead of asking for years of teaching experience with a range (0-5, 6-10, 11-15, etc.), it is better to ask for the specific number of years taught. The specific data can always be broken down into 0-5, 6-10, etc.; however, 0-5, 6-10 years can never be broken down into specific number of years.
  - Use plenty of white space on forms.

- **Instruments**
  - Gather data on specific constructs
  - Are usually developed by others

- **Norm-Referenced Testing**
  - Compares a student's performance on a number of skills against the performance of other students taking the test
  - Uses standardized procedures

- **Questionnaires**
  - Include questions about several issues
  - Are generally written by those gathering the information
- **Rating Scales**
  - Include Likert scales so that participants can indicate a range of scores, usually 1-5
  - Can be made up to fit the assessment, or can be commercially obtained
- **Surveys**
  - Are collections of questions that are asked of a large number of people
  - Can be done through the mail, in person, over the telephone, or with groups
  - *The Survey Kit* (Bourque, Fielder, Fink, & Litwin, 1995) contains nine books that provide information about conducting surveys. Bibliographic information is under the above-mentioned authors' names.
- **Additional Resources**
  - *How to Measure Performance and Use Tests* (Morris, Fitz-Gibbon, & Lindheim, 1987) provides information on using quantitative measures.

Qualitative Data

- **Used when:**
  - Detailed descriptions are desired
  - Goals of the program are not specific
  - Unanticipated outcomes are possible
  - A need exists to add depth and detail to quantitative findings
  - Understanding the dynamics is important
  - Individualized outcomes will result from the intervention
  - Quantitative measures to evaluate the outcomes do not exist

- **Case Studies**
  - Involve studying a limited number of situations in depth
  - Are used when unusual situations are likely to exist

- **Document Analysis**
  - Involves analyzing written materials, such as lesson plans, newspaper articles, minutes of meetings, etc.
  - Is less intrusive

- **Focus Groups**
  - *Focus Groups: A Practical Guide for Applied Research* (Krueger, 1994) provides information about how to conduct a focus group.
  - *Focus Group Interviews in Education and Psychology* (Vaughn, Schumm, & Sinagub, 1996) provides information about how to conduct a focus group.

- **Interviews**
  - Obtain more in-depth data for use in exploring an issue
  - Are used when probing beyond the initial response is desired
  - Rely on trained interviewers

- **Observations**
  - Provide the opportunity to experience the environment
  - Can use checklists can focus the observation
  - Should be long enough to gather meaningful data
  - Depend on trained observers for consistency
  - *How to Assess Program Implementation* (King, Morris, & Fitz-Gibbon, 1987) provides a chapter on conducting observations.

- **Videotapes**
  - Provide a means of capturing interactions for more in-depth analyses
  - Necessitate a high level of trust

- **Unobtrusive Measures**
  - Enable data to be gathered without anyone knowing about it
  - Could include gathering of student data, such as number of absences, grades, office referrals, number of books checked out of the library, etc. (Webb, Campbell, Schwartz, & Sechrest, 1966)

- **Additional Resources**
  - *How to Use Qualitative Methods in Evaluation* (Patton, 1987) provides in-depth information about qualitative methods.

Terms

Empowerment Evaluation (Fetterman, Kaftarian, & Wandersman, 1996)

- How can gathering the data involve participants in such a way that they will be empowered?
    - Example: Invite teachers who have been videotaped to analyze their videotapes and reflect on changes that they have made in their teaching.
    - Then, examine the teachers' analyses with the formal analyses of the videotapes.

Formative and Summative Evaluations

- Formative evaluations provide information to guide the project. They are conducted in an ongoing manner, and allow the directors to "steer the ship" based on feedback about what is working and what isn't working.
- Summative evaluations provide information at the end about the final result of the grant.

Qualitative Data Analysis - Two Types

- Content analysis - coding the data so that certain predetermined concepts go together
- Inductive analysis - allowing the patterns and themes to emerge out of the data

Triangulation

- Using multiple data collection strategies (Denzin, 1978)
- Strengthens the research design

Validity and Reliability

- Validity means that the instrument or method used actually provides the information that is needed.
- Reliability means that if the instrument were given a number of times, it would yield the same or similar results.

Sources of Instruments and Evaluation Strategies

- Do a search on ERIC or Dissertation Abstracts for the topic that you wish to explore. For example, a search on "Efficacy" will bring up research that has been done in that area. The abstract will usually name the instrument that was used, particularly if the search is done in Dissertation Abstracts. FirstSearch database, another one that is available, can be found at http://www.ref.oclc.org:2000. Descriptions of the databases can be found at http://www.oclc.org/oclc/fs/9085fs/descrev.htm. Other databases are Ovid and SilverPlatter.

- Do a search for instruments or topics on the Web by listing key words.

- Buros Tests and Measurements has a web site that lists numerous instruments (www.unl.edu/buros).

- The *Handbook of Tests and Measurements in Education and the Social Sciences* (Lester & Bishop, 1997) lists numerous resources that can be used.

- For instruments to measure student progress, check with the Assessment Department in your school district.

- The *Mental Measurements Yearbook* published by Buros [(402) 472-6203] provides numerous instruments.

- *Tests in Print*, also published by Buros, provides numerous instruments. It is available in many libraries.

- *Counseling Programs: A Guide to Evaluation* (Borders & Drury, 1992) provides suggested instruments and strategies for evaluating counseling programs.

- *Mathematics Programs: A Guide to Evaluation* (Bright, Uprichard, & Jetton, (1993) provides suggested instruments and strategies for evaluating mathematics programs.

- *Programs for At-Risk Students: A Guide to Evaluation* (O'Sullivan & Tennant, 1993) provides suggested instruments and strategies for evaluating programs for at-risk students.

- *Reading and Language Arts Programs: A Guide to Evaluation* (Olson & Miller, 1993) provides suggested instruments and strategies for evaluating reading and language arts programs.

- *Special Education Programs: A Guide to Evaluation* (Vallecorsa, deBettencourt, & Garriss, 1992) provides suggested instruments and strategies for evaluating special education programs.

- *Staff Development Programs: A Guide to Evaluation* (Mullins, 1994) provides suggested instruments and strategies for evaluating staff development programs.

Issues to Consider

Validity

- Does the instrument measure what it says it will measure?
- Does the instrument measure anything else?
- Does the instrument contain any language that might be offensive to certain groups?

Availability and Cost of the Instrument

- Is the instrument readily available?
- What is the cost of the instrument?

Reliability

- Would participants respond in the same way if they took the instrument today and next week?
- Are enough questions asked about a particular topic to make the instrument reliable?

Confidentiality of Participants

- How will confidentiality be maintained?
- Will participants be asked to sign an informed consent form?

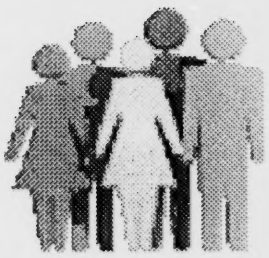
3. What types of data will indicate achievement of the intended results of the program? What are the variables?

4. What measures will indicate achievement of the intended results of the program?

Sources of Data / Participants

- Teachers
- Administrators
- Students
- Parents
- Community members
- Written information

5. Who will provide data to indicate success of the program?



Types of Research Designs

- A strong research design strengthens the final results.
- Random assignment of people to control and treatment groups strengthens the final results.
- If random assignment isn't possible, it is important to find a control group that is as similar to the treatment group as possible.
- Pretests are given before the treatment begins, mid-tests are given while the program is still being administered, and posttests are given at the end of the treatment.
- A series of tests, given before and after the program, is called "time series." Several tests given before the treatment begins can actually eliminate the need for a control group.
- Pretest / Posttest, Treatment / Control
    - Both treatment and control groups are administered a pretest and a posttest. Scores are compared to determine if the treatment group grew more than the control group.

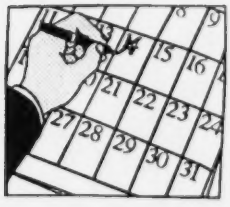
	Pretest	Posttest
Treatment Group		
Control Group		

- *How to Design a Program Evaluation* (Fitz-Gibbon & Morris, 1987) provides information about research designs.

6. What type(s) of research design(s) will be used?

	Pretest	Posttest
Treatment Group		
Control Group		

7. What timeline will be used for gathering and analyzing the data?



8. How will data be collected?

Data Analysis

Security

- First of all, keep all data in a safe place. If possible, make duplicates of the data to keep in a separate place.

Quantitative Data

- Small amounts of quantitative data can be analyzed by hand. Larger amounts can be analyzed using the Statistical Package for the Social Sciences (SPSS) (1-800-543-2185) or other statistical computer programs. Information about SPSS is available at the website, spss.com

- Detailed information about data analysis is provided in *How to Analyze Data* (Fitz-Gibbon & Morris, 1987). A section is also provided in *Evaluating School Programs: An Educator's Guide* (Sanders, 1992).

Qualitative Data

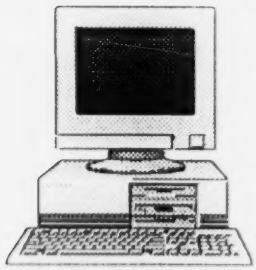
- Qualitative data can be analyzed either by hand or by computer. If data are analyzed by hand, the data are cut apart and categorized.

- If gathering qualitative data, it is important to budget for transcription of audio tapes. Generally, it takes 1 ½ to 2 times as long as the tape to transcribe them. For a one hour tape, it would be necessary to budget for 1 ½ to 2 hours of transcription time.

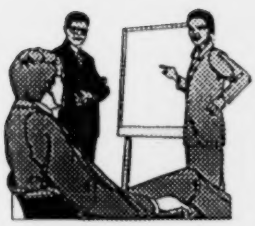
- A number of computer programs for analyzing qualitative data exist. *Computer Programs for Qualitative Data Analysis* (Weitzman & Miles, 1995) provides a detailed analysis of them. So many new programs have been developed recently that the book is in the process of being rewritten. NUD*IST and ATLAS/ti are perhaps the most powerful programs on the market. Both programs have the capability of converting data to SPSS files to drop into that program. Ethnograph is also good. TextSmart from SPSS is good for analyzing short responses.

- NUD*IST, ATLAS/ti, and Ethnograph can be ordered from Sage Publications Software at (805) 499-1325. The FAX is (805) 499-0871, and the address is 2455 Teller Road, Thousand Oaks, CA 91320. Sue at extension 7190 in the Scolari department is very helpful. Other programs that have developed recently are SphinxSurvey, Diction, winMAX, and Code-A-Text. Descriptions of the programs are available at the website, scolari.com.

9. How will data be analyzed?



10. How will the team interpret the data?

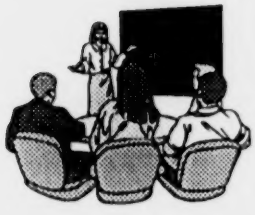


11. How will the team adjust the project based on findings?

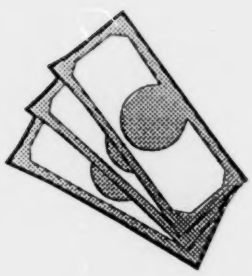
Communication of Findings

- Present an attractive report that matches the level of sophistication of the audience.
- Begin with the most important information.
- Highlight important points.
- Use simple words and avoid jargon.
- Provide the audience with time to talk about the findings and ask questions.
- Use graphs, tables, and figures to present the information visually.
- Provide the amount of information that the particular audience needs. For example, those who funded the program will expect a different type of report than community members, and those who participated in the program will have different expectations than the School Board.
- *Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations* (Torres, Preskill & Piontek, 1996) and *How to Communicate Evaluation Findings* (Morris, Fitz-Gibbon, & Freeman, 1987) provide additional information on this topic.

12. To whom will the results be communicated? How will they be communicated?



13. What will be the approximate cost of the evaluation?



Bibliography

Borders, L. D., & Drury, S. M. (1992). *Counseling programs: A guide to evaluation.* Newbury Park, CA: Corwin.

Bourque, L. B., & Fielder, E. P. (1995). *How to conduct self-administered and mail surveys.* Thousand Oaks, CA: Sage.

Anderson, S. B., & Ball, S. (1978). *The profession and practice of program evaluation.* San Francisco, CA: Jossey-Bass.

Brainard, E. A. (1996). *A hands-on guide to school program evaluation.* Bloomington, IN: Phi Delta Kappa Educational Foundation.

Bright, G. W., Uprichard, A. E., & Jetton, J. H. (1993). *Mathematics programs: A guide to evaluation.* Newbury Park, CA: Corwin.

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods.* Boston, MA: Allyn and Bacon.

Denzin, N. K. (1978). *The research act.* New York: McGraw-Hill.

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment & accountability.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *How to analyze survey data.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *How to ask survey questions.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *How to design surveys.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *How to report on surveys.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *How to sample in surveys.* Thousand Oaks, CA: Sage.

Fink, A. (1995). *The survey handbook.* Thousand Oaks, CA: Sage.

Fitz-Gibbon, C. T., & Morris, L. L. (1987). *How to analyze data.* Newbury Park, CA: Sage.

Fitz-Gibbon, C. T., & Morris, L. L. (1987). *How to design a program evaluation.* Newbury Park, CA: Sage.

Frey, S. H., & Oishi, S. M. (1995). *How to conduct interviews by telephone and in person.* Thousand Oaks, CA: Sage.

Henerson, M. E., Morris, L. L., & Fitz-Gibbon, C. T. (1987). *How to measure attitudes.* Newbury Park, CA: Sage.

Herman, J. L., Morris, L. L., & Fitz-Gibbon, C. T. (1987). *Evaluator's handbook.* Newbury Park, CA: Sage.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.

King, J. A., Morris, L. L., & Fitz-Gibbon, C. T. (1987). *How to assess program implementation.* Newbury Park, CA: Sage.

Krueger, R. A. (1994). *Focus groups: A practical guide for applied research* (2nd ed.). Thousand Oaks, CA: Sage.

Lester, P. E., & Bishop, L. K. (1997). *Handbook of tests and measurement in education and the social sciences.* Lancaster, PA: Technomic.

Litwin, M. S. (1995). *How to measure survey reliability and validity.* Thousand Oaks, CA: Sage.

Morris, L. L., Fitz-Gibbon, C. T., & Freeman, M. E. (1987). *How to communicate evaluation findings.* Newbury Park, CA: Sage.

Morris, L. L., Fitz-Gibbon, C. T., & Lindheim, E. (1987). *How to measure performance and use tests.* Newbury Park, CA: Sage.

Mullins, T. W. (1994). *Staff development programs: A guide to evaluation.* Thousand Oaks, CA: Corwin.

Olson, M. W., & Miller, S. D. (1993). *Reading and language arts programs: A guide to evaluation.* Newbury Park, CA: Corwin.

O'Sullivan, R. G. (Ed.). (1991). *Advances in program evaluation: A research annual* (Vol. 1, Part B). Greenwich, CT: JAI Press.

O'Sullivan, R. G., & Tennant, C. V. (1993). *Programs for at-risk students: A guide to evaluation.* Newbury Park, CA: Corwin.

Patton, M. Q. (1987). *How to use qualitative methods in evaluation.* Newbury Park, CA: Sage.

Pitz, G. F., & McKillip, J. (1984). *Decision analysis for program evaluators.* Beverly Hills, CA: Sage.

Posavac, E. J. & Carey, R. G. (1997). *Program evaluation: Methods and case studies* (5th ed.). Upper Saddle River, NY: Prentice-Hall.

Ralph, J., & Dwyer, M. C. (1988). *Making the case: Evidence of program effectiveness in schools and classrooms: Criteria and guidelines for the u. s. department of education's program effectiveness panel.* Washington, D. C.: Office of Educational Research and Improvement.

Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: A systematic approach* (5th ed.). Newbury Park, CA: Sage.

Sanders, J. R. (1992). *Evaluating school programs: An educator's guide.* Newbury Park, CA: Corwin.

Stake, R. E. (Ed.). (1991). *Advances in program evaluation: A research annual* (Vol. 1, Part A). Greenwich, CT: JAI Press.

Stecher, B. M., & Davis, W. A. (1987). *How to focus an evaluation.* Newbury Park, CA: Sage.

Stewart, D. W., & Shamdasani, P. N. (1990). *Focus groups: Theory and practice.* Newbury Park, CA: Sage.

Torres, R. T., Preskill, H. S., & Piontek, M. E. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations.* Thousand Oaks, CA: Sage.

Vallecorsa, A. L., deBettencourt, L. U., & Garriss, E. (1992). *Special education programs: A guide to evaluation.* Newbury Park, CA: Corwin.

Vaughn, S., Schumm, J. S., & Sinagub, J. (1996). *Focus group interview in education and psychology.* Thousand Oaks, CA: Sage.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.

Weitzman, E. B., & Miles, M. B. (1995). *Computer programs for qualitative data analysis.* Thousand Oaks, CA: Sage.

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). New York: Longman.

Intended Result: _____

Hypotheses	Type of Data	Variables	Participants	Measures	Timeline

Data Collection Procedures	Data Analysis Procedures	Interpretation Method	Plan to Adjust Project Based on Findings	Communication of Findings

Intended Result: _____

Hypotheses	Type of Data	Variables	Participants	Measures	Timeline

Data Collection Procedures	Data Analysis Procedures	Interpretation Method	Plan to Adjust Project Based on Findings	Communication of Findings